

أوامر تنزيل و إعدادات هادوب

تحديث النظام

```
$ sudo apt-get update
```

أولا

لتنزيل الجافا

```
$ sudo apt-get install default-jdk
```

```
$ java -version
```

ثانيا

تنزيل SSH من أجل اتصال الأجهزة عن بعد (خاصة بال multinode)

```
$ sudo apt-get install ssh
```

```
$ sudo apt-get install rsync
```

إنشاء وإعدادات SSH

```
$ ssh-keygen -t dsa -P "" -f ~/.ssh/id_dsa
```

```
$ cat ~/.ssh/id_dsa.pub >> ~/.ssh/authorized_keys
```

للفحص أنه يعمل

```
$ ssh localhost
```

ثالثا

لتنزيل هادوب (Hadoop)

```
$ wget http://mirrors.sonic.net/apache/hadoop/common/hadoop-2.6.5/hadoop-2.6.5.tar.gz
```

```
$ tar xvfz hadoop-2.6.5.tar.gz
```

ننقل هادوب إلى المسار **/usr/local/hadoop** الخاص بملفات directory

```
$ sudo mv hadoop-2.6.5 /usr/local/hadoop
```

```
$ update-alternatives --config java
```

رابعاً

إعدادت الملفات

يجب تعديل الملفات التالية لإكمال إعدادات هادوب

```
$ sudo gedit ~/.bashrc
```

نضيف في نهايته :

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export
HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"
```

```
$ source ~/.bashrc
```

ننتقل إلى المسار التالي من أجل إكمال تعديل الملفات الأخرى

```
$ cd /usr/local/hadoop/etc/hadoop
```

بعد الانتقال نبدء بالتعديل

```
$ sudo gedit hadoop-env.sh
```

```
export JAVA_HOME="/usr/lib/jvm/java-8-openjdk-amd64"
```

```
$ sudo gedit core-site.xml
```

```
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://localhost:9000</value>
  </property>
</configuration>
```

\$ sudo gedit yarn-site.xml

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-
services.mapreduce.shuffle.class</name>
    <value> org.apache.hadoop.mapred.ShuffleHandler</value>
  </property>
</configuration>
```

لتغيير اسم الملف الخاص بال mapred

\$ sudo cp mapred-site.xml.template mapred-site.xml

\$ sudo gedit mapred-site.xml

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

```
$ sudo gedit hdfs-site.xml
```

```
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/usr/local/hadoop/hadoop_data/hdfs/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/usr/local/hadoop/hadoop_store/hdfs/datanode</value>
  </property>
</configuration>
```

```
$cd
```

خامسا

إنشاء ملفات خاصة directory

```
$ mkdir -p /usr/local/hadoop/hadoop_data/hdfs/namenode
```

```
$ mkdir -p /usr/local/hadoop/hadoop_data/hdfs/datanode
```

```
$ sudo chown cloud:cloud -R /usr/local/hadoop
```

سادسا

تنسيق ملفات نظام هادوب

```
$ hdfs namenode -format
```

يستخدم لمرة واحدة

سابعاً

لتنشغيل هادوب

```
$ start-all.sh
```

أو من خلال الأمرين

```
$ start-dfs.sh
```

```
$ start-yarn.sh
```

لنتأكد من أنه يعمل أو لا

```
$ jps
```

ال ports الخاصة بالواجهات

```
http://localhost:8088/
```

```
http://localhost:50070/
```

```
http://localhost:50090/
```

```
http://localhost:50075/
```

الأوامر الخاصة بتجربة عداد الكلمات (word count)

بعد إنشاء ملف على سطح المكتب data وإدخال نص المراد عمل الماب ريديوس عليه :

نشغل الهادوب

```
$start-all.sh
```

ننتقل إلى المسار داخل الهادوب

```
$ cd /usr/local/hadoop
```

ننشئ ملف داخل directory وهو user

```
$ bin/hdfs dfs -mkdir /user
```

ننشئ داخل ال user ملف آخر ليكن cloud

```
$ bin/hdfs dfs -mkdir /user/cloud
```

ننقل ملف data من سطح المكتب الى ملف input داخل مسار /user/cloud

```
$ bin/hdfs dfs -put /home/cloud/Desktop/data input
```

الأمر الخاص بعمل ماب ريديوس على ملف input وإظهار النتائج داخل output

```
$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordcount input output
```

لإظهار النتائج من داخل output

```
$ bin/hdfs dfs -cat output/*
```